

Docket No. 218897US0

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Sergei V. MASHKO, et al.

GAU:

SERIAL NO: NEW APPLICATION

EXAMINER:

FILED: HEREWITH

FOR: EXPRESSION CONTROL SEQUENCE



REQUEST FOR PRIORITY

ASSISTANT COMMISSIONER FOR PATENTS
WASHINGTON, D.C. 20231

SIR:

- ☐ Full benefit of the filing date of U.S. Application Serial Number , filed , is claimed pursuant to the provisions of 35 U.S.C. §120.
- ☐ Full benefit of the filing date of U.S. Provisional Application Serial Number , filed , is claimed pursuant to the provisions of 35 U.S.C. §119(e).
- ☒ Applicants claim any right to priority from any earlier filed applications to which they may be entitled pursuant to the provisions of 35 U.S.C. §119, as noted below.

In the matter of the above-identified application for patent, notice is hereby given that the applicants claim as priority:

COUNTRY

Russia

APPLICATION NUMBER

2001104817

MONTH/DAY/YEAR

February 22, 2001

Certified copies of the corresponding Convention Application(s)

- ☒ are submitted herewith
- ☐ will be submitted prior to payment of the Final Fee
- ☐ were filed in prior application Serial No. filed
- ☐ were submitted to the International Bureau in PCT Application Number
Receipt of the certified copies by the International Bureau in a timely manner under PCT Rule 17.1(a) has been acknowledged as evidenced by the attached PCT/IB/304.
- ☐ (A) Application Serial No.(s) were filed in prior application Serial No. filed ; and
- ☐ (B) Application Serial No.(s)
- ☐ are submitted herewith
- ☐ will be submitted prior to payment of the Final Fee

Respectfully Submitted,

OBLON, SPIVAK, McCLELLAND,
MAIER & NEUSTADT, P.C.

Norman F. Oblon

Registration No. 24,618

C. Irvin McClelland

Registration Number 21,124



22850

Tel. (703) 413-3000
Fax. (703) 413-2220
(OSMMN 10/98)



РОССИЙСКОЕ АГЕНТСТВО ПО ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ
(РОСПАТЕНТ)

ФЕДЕРАЛЬНЫЙ ИНСТИТУТ ПРОМЫШЛЕННОЙ СОБСТВЕННОСТИ

рег. No 20/12-936

"18" декабря 2001 г.

U.S. PTO
10/068851



СПРАВКА

#9

Федеральный институт промышленной собственности Российского агентства по патентам и товарным знакам настоящим удостоверяет, что приложенные материалы являются точным воспроизведением первоначального описания, формулы и чертежей (если имеются) заявки на выдачу патента на изобретение № 2001104817, поданной в феврале месяце 22 дня 2001 года (22.02.2001).

Название изобретения

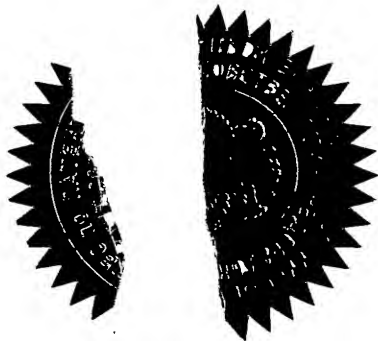
EXPRESSION CONTROL SEQUENCE

Заявитель

Закрытое акционерное общество
«Научно-исследовательский институт
Аджиномото-Генетика»

Действительный автор(ы)

МАШКО Сергей Владимирович
ЗИМЕНКОВ Данила Вадимович



Уполномоченный заверить копию
заявки на изобретение

А.Л. Журавлев
Заведующий отделом

EXPRESSION CONTROL SEQUENCE

BACKGROUND OF THE INVENTION

5 The present invention relates to microbiological industry, in particular, to the development of the new approach of the regulated gene expression in bacterial cells.

10 The inducing of expression of the cloned genes by the addition of the relatively simple and cheap chemical compounds is a very attractive idea for many biotechnological processes based on the exploiting of the recombinant bacteria, *E. coli*, in particular. Obviously, the natural L-amino acids are potentially very good candidates to be used as such inducers. But, 15 as far as inventors know, the regulatory regions of the tryptophanase genes are the unique natural systems induced by addition of tryptophan to the cultural medium [Landick R., Turnbough C.L., Yanofsky C. "Transcription attenuation"/ In: "*Escherichia coli* and *Salmonella*. 20 Cellular and molecular biology" (Second Edition, F.C.Neidhardt - Editor in Chief), (1996), pp. 1263-1286]. On the contrary, there are the several systems providing the decreasing of the controlled gene expression in case of excess of amino acids in cells (the system of the 25 *trp*-operon repressor - operator [Platt, T. "Regulation of gene expression in the tryptophan operon of *Escherichia coli*" / In: "The Operon" (Miller, J.H.,

Reznikoff, W.S. - Eds.), Cold Spring Harbor Laboratory (1978), pp. 263-302.], the attenuation of amino acid operon transcription [Landick R., Turnbough C.L., Yanofsky C. "Transcription attenuation"/ In:

5 "Escherichia coli and Salmonella. Cellular and molecular biology" (Second Edition, F.C.Neidhardt - Editor in Chief), (1996), pp. 1263-1286], in particular).

The molecular mechanism of transcription attenuation in amino acid operons is based on the possibility of the alternative mRNA secondary structures formation in the so-called "attenuator" region depending on the translation of the "leader" peptide of which gene is located upstream from the first structural gene of the operon. The coding part of the leader peptide gene is enriched by the codons of the sense amino acid (the codons of those amino acids whose biosynthesis are provided by the corresponding operons protein products: for the *trp*-operon there are the Trp-codons, for the *his*-operon - His-codons, for the *thr*-operon - Thr- and Ile-codons, etc).

The details of this well-established regulation of transcription are presented in the Fig. 1, using the attenuator regions of the *E. coli trp*- and *his*-operons for examples. As it is seen from the Fig. 1, the alternative mRNA secondary structures can be formed in case of transcription of the corresponding DNA fragments: the hairpins t1:t2 and t3:t4, or their

alternative - t2:t3, can be formed for *trp*-leader, analogously, h1:h2, h3:h4 and h5:h6, or h2:h3 and h4:h5 can be formed for *his*-leader. The hairpins t3:t4 and h5:h6 are the typical ρ -independent transcription terminators, so their formation during the elongation of transcription leads to termination in the attenuator regions of the corresponding operon preventing the expression of the structural genes of the operon.

As far as the process of mRNA secondary structure formation couples with the process of mRNA synthesis, so the hairpins t1:t2 and then t3:t4 are formed step by step (its alternative - the hairpin t2:t3 could not be formed because t1:t2 has been appeared earlier) in case of *trp*-attenuator when the leader peptide has not been translated. Analogously, the faster obtaining of the hairpins h1:h2, h3:h4 and then h5:h6 prevents the formation of their alternative - h2:h3 and h4:h5 in case of the synthesis of *his*-attenuator mRNA without translation of the corresponding leader peptide. As has been mentioned earlier, the formation of the hairpins t3:t4, as well as h5:h6 leads to termination in the corresponding attenuator regions. This situation could be realized during the transcription *in vitro* of the corresponding DNAs by the pure RNA polymerase without any translational factors in the reaction mixture, or *in vivo* in case of general amino acids starvation.

The much more complicate situation could be

occurred *in vivo* in case of translation of the leader peptide when the sense amino acid is in excess (more precisely, the corresponding charged tRNA is in excess), or in conditions of its deficiency in the bacterial cell.

5 It has been shown, that RNA polymerase initiating attenuator mRNA transcription, stops at the pause site in the region located immediately downstream from the hairpin 1:2 (probably, the formation of this hairpin is the essential, but not the adequate condition of such
10 pausing) [Chan, C.L., Wang, D., Landick, R. "Multiple interactions stabilize a single paused transcription intermediate in which hairpin to 3' end spacing distinguishes pause and termination pathway" / J. Mol. Biol. 268 (1997) 54-68]. The ribosome translating the N-
15 terminal part of the leader peptide, releases the RNA polymerase followed by continuation of the elongation of transcription, so the alternative hairpin of the mRNA - 2:3, can be formed. The following events depend on the intracellular concentration of the charged-tRNA(s) of
20 the sense amino acid, because the corresponding codons of the leader peptide gene have to be translated by the ribosome at this moment.

In case of the sense amino acid starvation the ribosome stalls at the sense codons and could not
25 disrupt the hairpin 2:3 while RNA polymerase synthesizes the downstream fragment of the mRNA which, in principle, could form the terminator hairpin (t3:t4 - for the *trp*-

attenuator and h5:h6 - for the *his*), but it does not fold due to the existence of the alternative hairpin (t2:t3 - for the *trp*, and the structure h2:h3, h4:h5 - for the *his*) followed by elongation of the transcription and synthesis of the mRNA of the operon structural genes.

In case of excess of the sense amino acid, the translation of the leader peptide occurs with the high efficiency and so, the ribosome initially disrupting the hairpin 1:2, disrupts the hairpin 2:3, as well, and stalls at the stop codon of the leader peptide. The later leads, finally, to formation of the terminator hairpin and to the termination of transcription in the attenuator region.

15 SUMMARY OF THE INVENTION

An object of the present invention is the creation of the new prokaryotic artificial regulatory system providing increase of the controlled gene expression as the result of increase of an intracellular concentration of the sense amino acid.

The present inventors have succeeded in creating a new artificial regulatory system depending on the intracellular amino acid concentration by the exploiting of the native regulatory mechanism based on the formation of the alternative secondary structures of mRNA in dependence on the efficiency of the leader peptide translation. The new system, unlike the natural

attenuators of amino acid operons, provides not decrease, but increase of the controlled gene expression in case of excess of the sense amino acid intracellular concentration. At the same time, the level of expression of genes under the control of the new expression control sequence is decreased in the conditions of the sense amino acid starvation. The well-known regulation of transcription, the attenuation of amino acid operon transcription, has been used as the progenitor of the new system, but the new system, unlike its progenitor, provides not the decrease, but increase of the controlled gene expression level in case of excess of the sense amino acid.

The present invention provides the expression control sequence as well as the expression control method and the production method using the expression control sequence, as mentioned below:

(1) An expression control sequence which controls expression of a target gene linked downstream of the expression control sequence depending on an intracellular concentration of an amino acid,

wherein in a bacterium which harbors a DNA construct comprising the expression control sequence, a promoter linked upstream of the expression control sequence and the target gene linked downstream of the expression control sequence, frequency of termination in the expression control sequence, of transcription

starting from the promoter is lowered by increase of an intracellular concentration of an amino acid, whereby expression of the target gene increases.

- (2) The expression control sequence according to (1), which comprises a region coding for a leader peptide comprising said amino acid and a ρ -independent terminator, wherein when translation of the leader peptide stops at codon of said amino acid in the course of the translation in case of starvation of the amino acid, a base pairing structure of the ρ -independent terminator is formed in a transcript of the expression control sequence, whereby the frequency of termination in the expression control sequence, of the transcription increases.
- (3) The expression control sequence according to (2), which comprises an odd number of not less than 3, of segments, wherein each of the segments can form a base pairing structure together with its adjacent segment, and wherein in the transcript of the expression control sequence, when a segment or segments other than terminal segments each form a base pairing structure with one of its two adjacent segments, the segment or segments each do not form a base pairing structure with the other of the two adjacent segments; a first segment at an upstream terminal overlaps with the region interacting with the ribosome translating the leader peptide; a second segment adjacent to the first segment forms a

base pairing structure with a third segment adjacent to the second segment in the course of the translation of the leader peptide; and a base pairing structure formed from the downstream terminal segment and its adjacent
5 segment is the base pairing structure of the ρ -independent terminator.

(4) The expression control sequence according to (3), wherein the first segment overlaps with codon of the amino acid in the leader peptide.

10 (5) The expression control sequence according to (3) or (4), wherein the number of the segments is 5.

(6) The expression control sequence according to any of (3) to (5), wherein the sequence of each segment or a part thereof and the sequence of the adjacent segment
15 or a part thereof constitute an inverted repeat sequence.

(7) The expression control sequence according to any of (2) to (6), wherein the ρ -independent terminator is capable of functioning in a bacterium belonging to the genus *Escherichia*.

20 (8) The expression control sequence according to (7), which comprises five segments a_1 to a_5 in order from an upstream side, wherein the segments a_1 and a_2 , and a coding region for the leader peptide are derived from a sequence of an attenuator of a tryptophan operon of
25 *Escherichia coli*, the segments a_4 and a_5 are derived from a sequence of an attenuator of a histidine operon of *Escherichia coli*, and the segment a_3 is derived from

a combination of the sequences of the attenuators of the tryptophan operon and the histidine operon.

(9) The expression control sequence according to (8), the leader peptide has been modified to contain not less than 2 of tryptophan residues.

(10) A method for controlling an expression of a target gene, comprising the steps of:

cultivating a bacterium harboring a DNA construct comprising the expression control sequence as defined in any of (1) to (9), a promoter linked upstream of the expression control sequence and the target gene linked downstream of the expression control sequence in a culture medium, and

changing an intracellular concentration of an amino acid on which expression control by the expression control sequence depend, to control expression of the target gene.

(11) A method for producing a target substance comprising the steps of cultivating a bacterium capable of producing the substance to produce the substance and collecting the substance,

wherein the bacterium harbors a DNA construct comprising the expression control sequence as defined in any of (1) to (9), a promoter linked upstream of the expression sequence and a target gene which has relationship to production of the target substance and is linked downstream of the expression control sequence,

and an intracellular concentration of an amino acid on which expression control by the expression control sequence depend, is changed to control expression of the target gene.

- 5 (12) The method according to (11), wherein the intracellular concentration of the amino acid is changed by synthesis or degradation of the amino acid by the bacterium.

According to the present invention, it is provided
10 an expression control sequence providing increase of the controlled gene expression by increasing an intracellular concentration of the sense amino acid. According to the expression control method of the present invention, expression of the target gene can be
15 increased by increasing the intracellular concentration of the sense amino acid. According to the production method of the present invention, the production amount of the target substance can be increased by increasing the intracellular concentration of the sense amino acid,
20 thus the target substance can be efficiently produced.

BRIEF EXPLANATION OF THE DRAWINGS

Fig. 1 shows the explanatory scheme of the structures and properties of the native attenuators and
25 the expression control sequence (artificial anti-attenuator) of the present invention.

Fig. 2 shows detailed structures of the native

attenuators and the artificial anti-attenuator. A and B represent the proposed "downstream" boarder of the parts of mRNA which are protected by the ribosome stalling at "sense" codons (A) and terminating at the stop codon (B) of the "leader" peptide, respectively. C represents the pause site of DNAs transcribing by E.coli RNA polymerase.

Fig. 3 shows the general scheme of construction of the artificial anti-attenuator. The cross-in-circle mark represents nucleotide changing in comparison to the native sequence. BI: *Bam*HI, Bg: *Bgl*II, NI: *Nde*I, XI; *Xba*I, XI*; *Xba*I(dam⁻).

Fig. 4 shows the scheme of construction of the plasmid in Example 1.

Fig. 5 shows the structure of the P_{tac} promoter from the plasmid pDR540. The sequences of the "upstream" (III) and "downstream" (IV) primers for PCR are dot-meshed.

Fig. 6 shows the structure of chemically synthesized an3:an4(an4:an5) fragment and the way the fragment was inserted in the cloning vector.

Fig. 7 shows the structure of the native attenuator region of the *E. coli trp*-operon. Leader peptide is shown in capital italic letters.

Fig. 8 shows the structure of the intermediate construction including fusion between RBS of bacteriophage T7 gene10 and the leader region of *trp* operon. Leader peptide is shown in capital italic

letters.

DETAILED DESCRIPTION OF THE INVENTION

<1> Expression Control Sequence

5 The expression control sequence of the present invention is an expression control sequence which controls expression of a target gene linked downstream of the expression control sequence depending on an intracellular concentration of an amino acid (sense
10 amino acid),

 wherein in a bacterium which harbors a DNA construct comprising the expression control sequence, a promoter linked upstream of the expression control sequence and the target gene linked downstream of the
15 expression control sequence, frequency of termination in the expression control sequence, of transcription starting from the promoter is lowered by increase of an intracellular concentration of an amino acid, whereby expression of the target gene increases.

20 The sense amino acid is not restricted provided that its aminoacyl tRNA can be synthesized and the synthesized aminoacyl tRNA can be used for translation of a protein in a bacterium for which the expression control sequence of the present invention is used.
25 Preferably, the sense amino acid is tryptophan, histidine, phenylalanine, threonine, leucine, isoleucine, or valine.

Examples of the target gene include chloramphenicolacetyltransferase gene (*cat*), amino acid operons, genes of which protein products are involved in the biosynthesis of amino acids, nucleosides and nucleotides, and genes encoding the foreign protein products.

Examples of the promoter include P_{tac} , and any other regulated and constitutive prokaryotic promoters.

Examples of the bacterium which harbors the DNA construct include bacteria belonging to the genres *Escherichia*, *Salmonella*, and *Serratia*.

The construction of the expression control sequence and the DNA construct and preparation of the bacterium harboring the DNA construct can be carried out according to standard gene engineering techniques (for example, see Molecular Cloning, 2nd Edition, Cold Spring Harbor Press (1989), Japanese Patent Application Laid-Open No. 2-207791 and the like).

The lowering of frequency of termination in the expression control sequence, of transcription starting from the promoter by increase of an intracellular concentration of an amino acid, whereby expression of the target gene increases, can be also determined according to the standard gene engineering techniques. With respect to the intracellular concentration of the sense amino acid, if a relationship between the intracellular and extracellular concentrations in the

bacterium is known, the intracellular concentration is not necessary to be directly measured and the intracellular concentration may be estimated based on the extracellular concentration such as a concentration in a medium. The frequency of termination in the expression control sequence, of transcription starting from the promoter is not necessary to be directly measured either and it is sufficient to determine increase of the target gene expression. The increase of the target gene expression can be determined by measuring an amount of the gene product of the target gene, or an activity of the gene product when the gene product has the activity.

An embodiment of the lowering of frequency of termination in the expression control sequence, of transcription starting from the promoter by increase of an intracellular concentration of an amino acid, whereby expression of the target gene increases, is exemplified by an embodiment in which the transcription starting from the promoter is terminated in the expression control sequence when the intracellular concentration of the sense amino acid is not higher than a certain level, and the transcription is elongated when the intracellular concentration of the sense amino acid is higher than the certain level, thereby expressing the target gene.

The expression control sequence of the present

invention preferably comprises a region coding for a leader peptide comprising the sense amino acid and a ρ -independent terminator, wherein when translation of the leader peptide stops at codon of the sense amino acid (sense codon) in the course of the translation in case of starvation of the sense amino acid, a base pairing structure of the ρ -independent terminator is formed in a transcript of the expression control sequence, whereby the frequency of termination in the expression control sequence, of the transcription increases.

The length of the leader peptide is usually 14 to 32 residues. The leader peptide usually contains 14 to 57%, preferably 30 to 45% of sense amino acid residues based on the total amino acid residues. As the proportion of the sense amino acid becomes large, control by the intracellular concentration of the sense amino acid becomes strict.

The ρ -independent terminator has a sequence which can form a base pairing structure (hairpin) and terminates the transcription when the base pairing structure is formed. The ρ -independent terminator is preferably one which is capable of functioning in a bacterium belonging to the genres *Escherichia*, *Salmonella*, or *Serratia*.

The means of allowing the base pairing structure of the ρ -independent terminator to be formed in the

transcript of the expression control sequence when translation of the leader peptide stops at the sense codons in the course of the translation, is exemplified by making the expression control sequence to comprise an

5 odd number of not less than 3, of segments, wherein each of the segments can form a base paring structure together with its adjacent segment, and wherein in the transcript of the expression control sequence, when a segment or segments other than terminal segments each

10 form a base paring structure with one of its two adjacent segments, the segment or segments each do not form a base paring structure with the other of the two adjacent segments; a first segment at an upstream terminal overlaps with the region interacting with the

15 ribosome translating the leader peptide; a second segment adjacent to the first segment forms a base paring structure with a third segment adjacent to the second segment in the course of the translation of the leader peptide; and a base paring structure formed from

20 the downstream terminal segment and its adjacent segment is the base paring structure of the ρ -independent terminator.

In this embodiment, it is necessary that the formation of the base paring structure between the first

25 segment and the second segment is blocked by stopping of the ribosome at the codon of the sense amino acid. Because of the mass of the ribosome, the ribosome covers

a region of from about 17 bp upstream from the codon of the sense amino acid where the ribosome stops, to about 13 bp downstream from the codon of the sense amino acid. The region interacting with the ribosome means such a region which is covered by the ribosome. If the first segment overlaps with the region interacting with the ribosome translating the leader peptide, it is predicted that the formation of the base pairing structure between the first segment and the second segment is sufficiently blocked by stopping of the ribosome at the codon of the sense amino acid. For example, if the codon of the sense amino acid exists immediately before the termination codon of the leader peptide and the distance from the codon of the sense amino acid to the starting point of the first segment is within about 13 bp, the formation of the base pairing structure between the first segment and the second segment can be blocked.

The first segment overlaps with the codon of the sense amino acid in the leader peptide.

In this embodiment, when the first segment is blocked by ribosome due to sense amino acid starvation, the base pairing structures between the second and the third, the forth and the fifth, ... are formed and the final base pairing structure functions as a terminator to terminate the transcription. On the other hand, when the sense amino acid is sufficiently provided and ribosome moves along mRNA to a stop codon at a

sufficient rate to follow the progress of the transcription, resulting in block of up to the second segment by ribosome, the base paring structures between the third and the forth, ... are formed to prevent
5 formation of a terminator.

In the expression control sequence of the present invention, it is preferred that a pause site for RNA polymerase exists in the region encoding the C-terminal region of the leader peptide, or downstream from the
10 region encoding the leader peptide, more preferably in a region from the second segment to the third segment. When no pause site exists, the expression of the target gene may not be sufficiently controlled if translation associated with transcription does not sufficiently
15 occur.

The number of the segments is, for example, 5.

The nucleotide sequence of each segment may be one which can form a base paring structure with its adjacent segment. For example, the sequence of each segment or a
20 part thereof and the sequence of the adjacent segment or a part thereof may constitute an inverted repeat sequence. The sequence constituting the inverted repeat may not be continuous. In other words, it may contain a part making no base pair within its sequence. In the
25 segment other than the terminal segments, it is sufficient that there is at least a partial overlap between a part constituting the inverted repeat sequence

with one of its adjacent segments and a part constituting the inverted repeat sequence with other of its adjacent segments.

5 An example of the expression control sequence is one which comprises five segments an1 to an5 in order from an upstream side, wherein the segments an1 and an2, and a coding region coding for the leader peptide are derived from a sequence of an attenuator of a tryptophan operon of *Escherichia coli*, the segments an4 and an5 are
10 derived from a sequence of an attenuator of a histidine operon of *Escherichia coli*, and the segment an3 is derived from a combination of the sequences of the attenuators of the tryptophan operon and the histidine operon.

15 The term "derived from" used herein means to have a sequence which is the same as or similar to the native sequence. The means of obtaining the sequence is not restricted. The sequence may be isolated from a biological material or chemically synthesized. The
20 similar sequence may be a sequence which has substitution, deletion or insertion of one or more nucleotides in the native sequence and which can form a base paring structure equivalent to that formed in the native sequence.

25 In the preferred example of the expression control sequence, the leader peptide is preferably one which has been modified to contain not less than 2 of tryptophan

residues. These tryptophan residues are preferably continuous.

The present invention is described with reference to the preferred example of the expression control sequence. The expression control sequence is
5 hereinafter referred to as an artificial anti-attenuator for convenience' sake.

The biological properties of the artificial anti-attenuator are schematically presented in the Fig 1. As
10 could be seen from the Fig. 1, in case of the sense amino acid starvation and the stalling of the ribosome at the corresponding codons of the leader peptide, the undisrupted hairpin an2:an3 of the anti-attenuator would conduce to formation of the hairpin an4:an5 which is the
15 part of the typical ρ -independent transcription terminator. On the other hand, the efficient translation of the leader peptide in case of excess of the sense amino acid would lead to the disruption of the hairpin an2:an3 of the artificial anti-attenuator
20 followed by the formation of the alternative hairpin an3:an4 that prevents the termination prior to transcription of the distal (downstream) genes, because the terminator hairpin an4:an5 could not be formed.

The construction of the artificial anti-attenuator
25 has been designed on the basis of two well-known native attenuators of the *E. coli trp*- and *his*-operons. It is used the native leader peptide gene of the *trp*-operon

(*trpL*) with two controlling Trp-codons in its structural part, as for the leader peptide of the artificial anti-attenuator. On the other hand, the process of the alternative mRNA secondary structure formation will take place in the 3'-untranslated region of the artificial anti-attenuator like in the corresponding region of the native *his*-attenuator.

The nucleotide sequences of the coding parts and the 3'-untranslated regions of two natural attenuators (*trp* and *his*) that were taken as the basic, as well as the sequence of the corresponding region of the artificial anti-attenuator are presented in the Fig. 2. As could be seen from the Fig. 2, the structures up to "+85" (C from the ATG of the coding part of the leader peptide has been marked as "+1") from the native *trp*-attenuator and from the artificial anti-attenuator coincide. So, it could be supposed that the processes of transcription elongation with pausing RNA polymerase in the position "+66 - +67", disruption of the hairpin an1:an2 by the ribosome translating of the leader peptide in case of Trp-starvation, as well as disruption in addition the hairpin an2:an3 in case of excess of Trp in the cell, would be occurred in the regulatory region of the artificial anti-attenuator just at the same manner as in the case of the native *trp*-attenuator. On the other hand, the distal part of the anti-attenuator significantly differs from the corresponding regions of

the *his*-attenuator, which has been used as the second progenitor.

The minimal alterations has been made in the region which could form the terminator hairpin an4:an5. Nevertheless, as in the native *his*-attenuator, the possible formation of the alternative hairpin an3:an4 has been provided in the artificial anti-attenuator (the secondary structure of this region is homologous to h4:h5 of the native *his*-attenuator) whose formation could prevent the termination of transcription in the artificial anti-attenuator. In comparison with the native *his*-attenuator, in the anti-attenuator there is no DNA fragment providing formation of the h3:h4 hairpin followed by the changing of the biological properties thereof: formation of the p-independent transcription terminator takes place in case of the sense amino acid starvation instead of its increased intracellular concentration.

The results of confirmation of the properties of the artificial anti-attenuator are described below.

The artificial anti-attenuator has been synthesized using the standard gene engineering techniques (including the PCR-driven amplification of the native fragment of the *trp*-attenuator, chemical synthesis of the oligonucleotides etc.) as schematically presented in the Fig. 3. Two different types of the artificial anti-attenuators have been obtained. The

native ribosome binding site (RBS) of *trpL* has been used for providing the translation initiation of the leader peptide in the artificial anti-attenuator of the first type - anti-attenuator-I. More efficient RBS of phage T7 gene10 has been inserted in the 5'-region of the leader peptide gene in the second - anti-attenuator-II. The both artificial anti-attenuators have been cloned in the vector plasmid downstream the high-efficient promoter P_{tac} and upstream of the structural part of *cat* gene with its own RBS. The *cat* gene encoding the chloramphenicolacetyltransferase (CAT), has been used as a reporter and its level of expression could give the information concerning of the function efficiency of the created regulatory elements in dependence on the intracellular concentration of the sense amino acid - Trp.

Moreover, the several control recombinant plasmids have been constructed on the basis of the same vector. The first plasmid carries the native attenuator of the *trp*-operon - *trpL*. The second - the potential transcription terminator - the 3'-terminal DNA fragment of the anti-attenuators providing the formation of the terminator hairpin an4:an5. The third plasmid carries the longer 3'-terminal DNA fragment of the anti-attenuators that could not provide the formation of the terminator hairpin an4:an5, because the potential alternative hairpin an3:an4 has to be formed first

during the mRNA synthesis.

It is known from the literature, that the native *trpL* of the *E. coli* *trp*-operon (unlike the attenuators of other amino acid operons) is rather weak attenuator:

5 "The presence or absence of Trp in the growth medium does not normally affect readthrough of the *trp* attenuator; the 10-fold change in readthrough occurs only upon extreme Trp starvation in mutant bacteria or transiently upon transfer of bacteria from a Trp-

10 containing to a Trp-free medium" [Landick R., Turnbough C.L., Yanofsky C. "Transcription attenuation"/ In: "*Escherichia coli* and *Salmonella*. Cellular and molecular biology" (Second Edition, F. C. Neidhardt - Editor in Chief), (1996), p. 1263-1286]. The later could be,

15 probably, due to the presence of only tandem codons of the sense amino acid in the structural part of the corresponding leader peptide. Since the coding parts of the leader peptides of the artificial anti-attenuators and the native *trpL* are identical, the special model

20 system has to be used for testing the regulatory properties of the new systems. The determination of the enzymatic activity of the reporter CAT has been provided in the *trp*⁻ cells carrying the recombinant plasmids of interest, and growing in condition of Trp-starvation

25 followed by addition of Trp in the cultural medium, if necessary. The obtained results are presented in the Table 1.

Table 1. Determination of CAT activity in the strains carrying the recombinant plasmids with the tested regulatory elements.

Plasmid name	Proposed properties	CAT activity in conditions of Trp:	
		Starvation	Excess
PML-P _{tac} -an4:an5-cat	"terminator"	3 ± 1	3 ± 1
PML-P _{tac} -an3:an4(an4:an5)-cat	"antiterminator"	60 ± 6	58 ± 6
PML-P _{tac} -trpL-cat	Native trpL	17 ± 3	9 ± 3
PML-P _{tac} -anti_att-I-cat	"anti-attenuator-I"	22 ± 3	30 ± 3
pML-P _{tac} -anti_att-II-cat	"anti-attenuator-II"	18 ± 3	51 ± 5

As could be seen from the Table 1, the level of CAT activity does not depend on the addition of the Trp to the cells carrying the plasmids without the coding part of the *trpL*. Moreover, the theoretical difference in the achieved level of CAT activities depending on the process of the alternative mRNA secondary structure formation, could be evaluated on the basis of the obtained results for the control plasmids encoding the "terminator" and "antiterminator" hairpins. The 20-th fold of transcription increase could be achieved in case of formation of the "antiterminator" structure in comparison with the ρ -independent transcription termination in the 3'-part of the artificial anti-attenuators.

As could be suspected, the determined level of CAT

activity has been higher in case of exploiting of the plasmid with the native *trp*-attenuator: the increased level has been achieved in case of Trp-derived starvation, than after addition of Trp to the growing plasmid-carrier bacteria. The concrete achieved level of CAT activity (17 units) could not be compared with the maximum level (60 units). That is because, at first, the Trp-derived starvation in case of bacterial growing, does not guarantee the maximal efficiency of the *trpL* readthrough transcription. The second, the 5'-untranslated region of *cat* gene located immediately upstream its own RBS, in case of exploiting of the native *trpL* and other tested constructions are different, so the efficiency of CAT translation initiation could differ, as well.

The main result has been obtained for the plasmids carrying the artificial anti-attenuators. As could be seen from the Table 1, the increased level of CAT activity after addition of Trp to the growing bacteria could be seen for the both plasmid-carrier bacterial strains. Moreover, exploiting of the high-efficient RBS of phage T7 *gene10* for translation initiation of the leader peptide leads to achievement of the level of CAT accumulation which is closed to the theoretical maximum (51 units in comparison with 60 units). The later could be explained by the fact, that it is necessary to improve the RBS of the leader peptide for achievement of

the full transcription-translation coupling essential for realization of molecular mechanism of alternative mRNA secondary structure formation (typical for the native attenuation transcription in prokaryotic amino acid operons) in case of transcription of the artificial anti-attenuator from the rather strong promoter P_{tac} (which is stronger than the native promoter of *trp*-operon).

The following conclusions could be done on the basis of the obtained results:

1. The process of alternative mRNA secondary structure formation could be really exploited for regulation of transcription: the differences in the expression level could be achieved up to 20-fold due to exploiting of the artificial "his-like" tails.
2. The obtained artificial "anti-attenuator" systems were functionally active:
 - 2.1. The presence of the coding part of the native *trp*-leader peptide could stimulate the process of mRNA secondary structure folding depended on the intracellular Trp-concentration.
 - 2.2. The best "activator"-effect in excess of Trp has been found in case of exploiting of P_{tac} -promoter for transcription in combination with the optimized translation initiation of the leader peptide (exploiting of the phage T7 S10 RBS instead the native RBS of *trpL*).

3. It could be supposed that the developed system could be used as a basis of creation of the really inducible regulatory element: it could be induced by addition of the excess of Trp in the cultural medium after increasing of the quantity of the sense (Trp) amino acid codons in the structural part of the leader peptide of anti-attenuator region.

<2> Method for Controlling Target Gene Expression

The expression control method of the present invention is a method for controlling an expression of a target gene, comprising the steps of:

cultivating a bacterium harboring a DNA construct comprising the expression control sequence of the present invention, a promoter linked upstream of the expression control sequence and the target gene linked downstream of the expression control sequence in a culture medium, and

changing an intracellular concentration of an amino acid (sense amino acid) on which expression control by the expression control sequence depend, to control expression of the target gene.

The DNA construct and the bacterium harboring the DNA construct may be as described above with respect to the expression control sequence.

The culture conditions are not restricted provided that the bacterium can survive. The conditions are

properly selected depending on the bacterium.

The method of changing an intracellular concentration of the sense amino acid, is exemplified by a method of changing a concentration of the sense amino acid in the medium in which the bacterium is cultured, a method of changing an amount of synthesis or degradation of the sense amino acid in cells. The change of the target gene expression can be determined by measuring an amount of the gene product of the target gene, or an activity of the gene product when the gene product has the activity.

<3> Method for Producing Target Substance

The production method of the present invention is a method for producing a target substance comprising the steps of cultivating a bacterium capable of producing the substance to produce the substance and collecting the substance,

wherein the bacterium harbors a DNA construct comprising the expression control sequence of the present invention, a promoter linked upstream of the expression sequence and a target gene which has relationship to production of the target substance and is linked downstream of the expression control sequence, and an intracellular concentration of an amino acid (sense amino acid) on which expression control by the expression control sequence depend, is changed to

control expression of the target gene.

Examples of the target substance include CAT, and other prokaryotic enzymes, foreign protein products, amino acids, nucleotides and nucleosides, vitamins, and other biological active substances.

Examples of the bacterium capable of producing the target substance include bacteria belonging to the genres *Escherichia*, *Salmonella*, and *Serratia*.

The culture conditions are not restricted provided that the bacterium capable of producing the target substance can produce the target substance. The conditions are properly selected depending on the bacterium.

Cultivation is usually carried out under an aerobic condition for 10 to 50 hours. The cultivation temperature is usually controlled at 28 to 37°C, and pH is usually controlled at 6.6 to 7.4 during cultivation. Inorganic or organic, acidic or alkaline substances as well as ammonia gas or the like can be used for pH adjustment.

The medium may be an ordinary medium containing a carbon source, a nitrogen source, organic ions and optionally other organic components.

As the carbon source, it is possible to use sugars such as glucose, lactose, galactose, fructose, sucrose or starch hydrolysate; alcohols such as glycerol or sorbitol; or organic acids such as fumaric acid, citric

acid or succinic acid.

As the nitrogen source, it is possible to use inorganic ammonium salts such as ammonium sulfate, ammonium chloride or ammonium phosphate; organic
5 nitrogen such as soybean hydrolysate; ammonia gas; or aqueous ammonia.

It is desirable to allow required substances such as vitamin B₁ or yeast extract to be contained in appropriate amounts as organic trace nutrients. Other
10 than the above, potassium phosphate, magnesium sulfate, iron ion, manganese ion and the like are added in small amounts, if necessary.

Collection of the target substance from a culture such as cells and a culture medium can be usually
15 carried out by combining an ion exchange resin method, a precipitation method and other known methods.

The DNA construct may be as described above with respect to the expression control sequence provided that a gene which has relationship to production of the
20 target substance is used as the target gene. Examples of the target gene which has relationship to production of the target substance include biosynthesis genes for the target substance, genes for production of energy and related substances such as intermediates used for
25 biosynthesis of the target substance, regulatory genes therefor and the like. Specific examples thereof include L-tryptophan biosynthetic genes, L-serine

biosynthetic genes (for L-tryptophan biosynthesis, in particular), pntAB genes, genes of H⁺-ATPase.

In particular, if L-tryptophan is used as the sense amino acid and an L-serine biosynthetic gene, as the target gene, is linked downstream from the expression control sequence of the present invention in an L-tryptophan-producing bacterium, the expression of the L-serine biosynthetic gene increases when an intracellular accumulation amount of L-tryptophan increases. The time when the intracellular accumulation amount of L-tryptophan increases is also a time when L-serine which is one of substrates for L-tryptophan biosynthesis decreases. Therefore, the expression of the L-serine biosynthetic gene can be increased only when L-serine decreases. On the other hand, the L-serine biosynthetic gene can not be always, highly expressed in the L-tryptophan-producing bacterium, because a high concentration of L-serine is harmful to growth of cells. It is accordingly appreciated that the expression control sequence of the present invention is very useful when it is intended to increase the expression of the L-serine biosynthetic gene in the L-tryptophan-producing bacterium.

The bacterium which is capable of producing the target substance and harbors the DNA construct may be obtained by allowing a bacterium capable of producing the target substance to harbor the DNA construct or

conferring the ability to produce the target substance on a bacterium harboring the DNA construct. The allowing the bacterium to harbor the DNA construct can be carried out according to standard gene engineering techniques. The conferring of the ability to produce the target substance can be carried out according to the known method. For example, when the target substance is CAT, a method of introduction of a DNA encoding the chloramphenicol acetyl transferase can be used.

The method of changing an intracellular concentration of the sense amino acid, is exemplified by a method of changing a concentration of the sense amino acid in the medium in which the bacterium is cultured, a method of changing an amount of synthesis or degradation of the sense amino acid in cells. The method of changing an amount of synthesis or degradation of the sense amino acid in cells is preferable, because the production of the target substance can be couple with the production of an intermediate and the like for the target substance production. The change of the target gene expression can be determined by measuring an amount of the gene product of the target gene, or an activity of the gene product when the gene product has the activity.

EXAMPLES

Example 1. Construction of the recombinant plasmids carrying the native attenuator and the artificial anti-attenuators and their fragments

5 1. Construction of the vector plasmid pML-P_{tac}-ter_{thrL}-cat

 The plasmid vector pML-P_{tac}-ter_{thrL}-cat carrying ColE1-like replicon, Ap^R-gene as a selective marker, the tac promoter (P_{tac}) (Russel D.R., Bennett G., Gene 20
10 (1982) 231), synthetic ρ-independent transcription terminator of the leader peptide of the *E. coli* thr-operon (ter_{thrL}) and the structural part of cat-gene downstream of P_{tac}, has been constructed in the following manner. The gene cat transcribing from P_{tac}, has been
15 used as a reporter in the further experiments.

1.1. Construction of the pML-pp-vector

 Two plasmids have been used as the progenitors for construction of pML-pp-vector. The first one was our
20 previously described plasmid pML24 (Trukhan et al., Biotechnologiya (in Russian) 4, No. 3 (1988) 325-334). The second plasmid was the commercially available (MBI "Fermentas", Lithuania) vector pUC57 (GenBank/EMBL accession number Y14837). The scheme of construction of
25 pML-pp-vector based on the standard gene engineering procedures (Sambrook et al., "Molecular cloning.

Laboratory manual". (1989) Second Edition, Cold Spring Harbor Laboratory Press), is presented in the Figure 4.

1.2. Insertion of the chemically synthesized *ter_thrL* in the plasmid pML-pp-vector

The synthetic ρ -independent transcription terminator of the *E. coli thr*-operon leader peptide has been constructed due to the annealing of two chemically synthesized oligonucleotides with the following

structures:

I - 5'-ctagaaagcttaacacagaaaaaagcccgcacctgacagtgcgggcttt
ttttttcgaccactgcagg→3' (SEQ ID NO: 4), and

II - 5'-gatccctgcagtggtcgaaaaaaaaagcccgccactgtcaggtgcgggc
ttttttctgtgttaagcttt→3' (SEQ ID NO: 5).

The double-stranded DNA fragment with the single-stranded "cohesive" ends typical for obtaining after the *XbaI*- and *BamHI*-derived DNA treatment, has been constructed after the annealing of the synthetic oligonucleotides.

This fragment has been phosphorylated by T4 polynucleotidekinase according to the standard procedure and cloned in the plasmid pML-pp-vector cleaved by *XbaI* and *BamHI*. The correlation between the desired and obtained plasmid structure has been established due to restriction analysis and DNA sequencing of the inserted fragment.

1.3. Molecular cloning of P_{tac}

The PCR-driven DNA amplification has been provided for molecular cloning of the hybrid promoter (P_{tac}). Two oligonucleotides have been chemically synthesized for this purpose:

III - 5'-gcttaggtaccctcccatccccctgttgac→3' (SEQ ID NO: 6), and

IV - 5'-ctgtttctagatcctgtgtgaaattgttatccgc→3' (SEQ ID NO: 7);

and the commercially available plasmid pDR540 (Pharmacia, Sweden) carrying P_{tac} -promoter, has been used as a template for the PCR.

These oligonucleotides carried the sequences upstream (III) and downstream (IV) of the promoter as presented in the Figure 5. Moreover, they carried the sequences recognized by several (*KpnI* and *XbaI*) restriction endonucleases (see, Figure 5) for convenience of the following molecular cloning. The amplified DNA fragment has been treated by *KpnI* and *XbaI* followed by its molecular cloning in the previously obtained (see, item 1.2) vector plasmid cleaved by the same restrictases. The selected recombinant plasmid has been named as pML- P_{tac} →ter_thrL→cat and used as a vector in the further experiments.

25

2. Construction of the plasmids with the native and artificial transcription regulatory elements

The above described plasmid pML-P_{tac}→ter_thrL→cat has been used as a vector for creation of all plasmids of interest. Moreover, the set of oligonucleotides has been chemically synthesized:

- 5 olig1: 5'-cagagctctagaagttcacgtaaaaagggtatcgac-3' (SEQ ID NO: 8);
- olig2: 5'-gtatcgcatatgaaagcaattttcgtactgaaagg-3' (SEQ ID NO: 9);
- 10 olig3: 5'-gtctgagatctagtatctgattgctttacgcatgggtg-3' (SEQ ID NO: 10);
- olig4: 5'-atcataggatcctaattttgttcaaaaaaagcccgctcatt-3' (SEQ ID NO: 11);
- olig5: 5'-cgactgtctagaacggtacagaaagcccccggcagat-3' (SEQ ID NO: 12);
- 15 cat3': 5'-agctcaccgtctttcattgccatacgg-3' (SEQ ID NO: 13);
- cr5': 5'-acatgcggtaccgatcccgcgaaattaatac-3' (SEQ ID NO: 14).

- 20 These oligonucleotides have been used as the primers for the PCR-driven DNA amplification, as described below. The other set of oligonucleotides:
- olig6: 5'-cagagctctagaagatctgcccgactgcgtacaacggtacagaaagcccccggcagatcacctgc-3' (SEQ ID NO: 15);
- 25 olig7: 5'-cggggggcttttttattgcgcggttgataacgggatccagcgta-3' (SEQ ID NO: 16);
- olig8: 5'-tacgctggatcccgttatcaaccgcgcaataaaaaagcccccggcaggatctgccgggggctt-3' (SEQ ID NO: 17);

olig9: 5'-tctgtaccgttgtacgcagtcgggcagatcttctagagctctg-3'
(SEQ ID NO: 18),

has been used directly in the process of reconstruction
of the 3'-terminal part of the new artificial anti-
5 attenuator (see, below).

2.1. Construction of the plasmid pML-P_{tac}- an3:an4(an4:an5)-cat

This plasmid has been constructed on the basis of
10 pML-P_{tac}→ter_{thrL}→cat due to insertion of the double-
stranded DNA fragment created from the chemically
synthesized oligonucleotides (olig6, olig7, olig8 and
olig9), instead of ter_{thrL} between promoter P_{tac} and the
structural part of cat-gene (see, Figure 6). For this
15 purpose initially 650 ng of olig7 and 650 ng of olig9
have been phosphorylated by T4 polynucleotidekinase
("MBI Fermentas", Lithuania) according to the
recommended protocol. Two mixtures contained 430 ng of
olig6 plus 650 ng of phosphorylated olig9 and 430 ng of
20 olig8 plus 650 ng of phosphorylated olig7 in 30 µl of
the "Y+/Tango" buffer ("MBI Fermentas", Lithuania) has
been heated under 100°C for 5 min and then annealed at
75°C for 5 min. After that they were mixed together,
heated under 60°C for 5 min and annealed at 20°C for 10
25 min. This was followed by addition of 5 units of T4 DNA
ligase ("MBI Fermentas", Lithuania) and 0.5 µl of 100 mM
ATP and incubation at 22°C for 4 hours and overnight

incubation at +4°C. The mixture was heated at 68°C for 10 min and loaded on gel-electrophoresis. The well-seen DNA band was isolated using Low-Melting-Point Agarose technique. The obtained double-stranded DNA fragment (108 bp in length) has been treated by *Xba*I ("MBI Fermentas", Lithuania) as recommended by producer. 360 ng of this fragment has been ligated to 50 ng of vector pML-P_{tac}→ter_thrL→cat which has been prepared from *E. coli* (*dam*⁻) strain and cleaved by *Xba*I. In all cases when the vector plasmid pML-P_{tac}→ter_thrL→cat and the recombinant plasmids obtained on its basis have to be cleaved by the restrictase *Xba*I, the plasmid DNA has to be provided from the *E. coli* (*dam*⁻) strain because the *Xba*I-restriction site in these plasmids is overlapped with the GATC-sequence, which is a target of the Dam-driven DNA modification, and so the plasmid DNA purified from the *E. coli* (*dam*⁺) strain could not be cleaved by *Xba*I. Ligation has been performed with 3 u of T4 DNA Ligase at +4°C during the night. The obtained mixture was treated with *Bam*HI ("MBI Fermentas", Lithuania) according to the standard protocol. At the last step this DNA mixture has been diluted to 60 µl volume of T4 DNA Ligase buffer and has been treated with 5 units of T4 DNA Ligase overnight at +4°C. The resulting mixture has been transformed to the strain HB101 and colonies have been screened for desired construction. So, the plasmid pML-P_{tac}-an3:an4(an4:an5)-cat has been obtained

and its structure has been confirmed by restriction analysis and DNA sequencing according to the standard Sanger's procedure. We suspect, that the obtained plasmid carries the DNA fragment which is transcribed, could provide the formation of anti-terminator hairpin an3:an4 (the terminator hairpin an4:an5 could not form because an3:an4 being synthesized earlier).

2.2. Construction of the plasmid pML-P_{tac}-an4:an5-cat

The above described plasmid pML-P_{tac}-an3:an4(an4:an5)-cat has been as a progenitor for construction the next recombinant DNA: it has been as a template for PCR-driven DNA amplification of the fragment encoding the last, an4:an5, hairpin of the new regulatory region. For this PCR the oligonucleotides: olig5 and cat3' have been used as the primers (see, Figure 6). The first of them corresponds to the beginning of those part of the earlier cloned fragment that encodes the hairpin an4:an5, but it has, in addition, the nucleotides recognized by XbaI near the 5'-terminus (see, Figure 6). The second oligonucleotide, cat3', corresponds to the fragment of the coding part of the cat-gene (position +219 - +245, if A from ATG-initiation codon of CAT is numbered as "+1") (see, Figure 6). The double-stranded DNA fragment obtained due to PCR, has been treated by XbaI, ligated with the vector plasmid pML-P_{tac}→ter_thrL→cat cleaved by the same

restrictase, followed by *Bam*HI-treatment and
 recyclization of the product by T4 DNA ligase. So, the
 plasmid of interest, named pML-P_{tac}-an4:an5-cat, has been
 obtained.

5

2.3. Construction of the plasmid pML-P_{tac}-trpL-cat

For construction of plasmid carrying the native
 gene of the *E. coli* trp-operon leader peptide (gene
trpL) under the transcriptional control of P_{tac}-promoter,
 10 the chromosomal DNA from the strain *E. coli* MG1655, the
 whole genome sequence of which had been determined, has
 been used as a template for the PCR. The
 oligonucleotides olig1 and olig4 corresponding to 5'-
 and 3'-terminal parts of the *trpL*-gene (see, Figure 7),
 15 have been used as the primers for DNA amplification. As
 could be seen from the Figure 7, these primers carry in
 addition the flanking *Xba*I and *Bam*HI (in olig1 and olig4,
 correspondingly) recognition sites for the convenience
 of the following manipulation. The double-stranded DNA
 20 fragment 175 bp in length, has been treated by *Xba*I and
*Bam*HI followed by its cloning in the vector plasmid pML-
 P_{tac}→ter_thrL→cat cleaved by the same restrictases. The
 obtained plasmid carrying the native *trpL*-gene instead
 of ter_thrL in the vector plasmid, was named as pML-P_{tac}-
 25 trpL-cat.

2.4. Construction of the plasmid pML-P_{tac}-anti_att-I-cat

The above described plasmid pML-P_{tac}-trpL-cat has been used as a template in PCR for creation of the next recombinant DNA. In this procedure the previously described oligonucleotide, olig1, as well as the olig3 has been used as the primers. The olig3 corresponds to the central part of the native trpL-gene and carries in addition the BglII-recognition site at its 5'-terminus (see, Figure 7). After the PCR-driven DNA amplification the obtained double-stranded fragment 133 bp in length has been treated by XbaI, ligated with the plasmid pML-P_{tac}-an3:an4(an4:an5)-cat cleaved by the same restrictase, followed by hydrolysis of the product by BglII and recyclization by T4 DNA ligase. The obtained plasmid carrying the artificial anti-attenuator between P_{tac}-promoter and the structural part of the cat-gene, was named as pML-P_{tac}-anti_att-I-cat.

2.5. Construction of the plasmid pML-P_{tac}-anti_att-II-cat

The construction of the recombinant plasmid carrying the artificial anti-attenuator with the high-efficient ribosome binding site (RBS) of the phage T7 gene10 upstream the coding part of the native leader peptide of the *E. coli* trp-operon, has been provided in the following manner. Initially, the double-stranded DNA fragment has been obtained due to the PCR using olig2 (see, Figure 7) and olig3 as the primers and DNA of the plasmid pML-P_{tac}-trpL-cat as a template. So, the

restriction site for *NdeI* has been reconstructed upstream of ATG-initiating codon of the leader peptide (the nucleotides of ATG-codon are the part of the sequence CATATG recognized by *NdeI*). The obtained DNA fragment cleaved by *NdeI* has been ligated with the commercially available ("Novagen", USA) plasmid vector pET-22b(+) treated by *NdeI*. The plasmid pET-22b(+) carries RBS of T7 gene10 between *XbaI* and *NdeI* restriction sites. The product of this ligation has been used as a template for the PCR at the next stage of construction. The new oligonucleotide cr5' (see, Figure 8) as well as the previously exploited olig3 have been used as the primers for this PCR. The obtained double-stranded DNA fragment 210 bp in length, has been treated by *XbaI* and *BglIII* and cloned in the plasmid pML-P_{tac}-an3:an4(an4:an5)-cat according to the protocol exploited for construction of pML-P_{tac}-anti_att-I-cat. So, the new plasmid named as pML-P_{tac}-anti_att-II-cat carrying the artificial anti-attenuator with RBS of phage T7 gene10 in the 5'-untranslated region of the *trp* leader peptide gene, has been obtained. The structures of all plasmids carrying the artificial transcriptional regulatory elements, have been confirmed by the restriction analysis and sequencing according to the standard Sanger's procedure.

Example 2. The detection of the accumulation levels of Cat protein in strains carrying the recombinant plasmids with the native *trpL*, the artificial anti-attenuators and their fragments.

5

The previously described plasmids (see, Example 1): pML-Ptac-an4:an5-cat, pML-Ptac-an3:an4(an4:an5)-cat, pML-Ptac-*trpL*-cat, pML-Ptac-anti_att-I-cat, pML-Ptac-anti_att-II-cat have been introduced into the strains

10 *E.coli* TG1 (*supE*, *hsd*, *thi*, $\Delta(lac-proAB)$, $F'[traD36, proAB^+, lacI^Q, lacZ\Delta M15]$) and *E.coli* B7248(*trpB*:*Tn10*, *Str^R*) according to the standard experimental protocols with the selection of the plasmid-carrier cells on the medium with the ampicillin (100 μ g/ml) addition

15 (Sambrook et al., "Molecular cloning. Laboratory manual".(1989) Second Edition, Cold Spring Harbor Laboratory Press). The obtained cell cultures were grown in the tubes with the liquid medium at 37°C with good aeration. As for the cultivation medium, the L-broth

20 with the ampicillin addition was used for the TG1-driven plasmid-carrier strains and the minimal M9-media with ampicillin(100 μ g/ml), thiamine (5 μ g/ml) and tryptophan (10 μ g/ml) for the strains constructed on the basis of *E.coli* B7248. The overnight B7248-driven cultures were

25 diluted in 50 times with the same cultural media and the cultivation had been continued for 2-4 hours until the optical density at 600 nm being 1($OD_{600}=1$). Each of the

culture media was divided into two parts and tryptophan (200 µg/ml) was added to the one of two portions. The cultivation was continued for 1 hour, then cells were collected by centrifugation, washed with physiological solution and resuspended in 1/10 of the initial volume of potassium-phosphate buffer. Then probes were sonicated, debris was harvested by centrifugation at 4°C. The protein concentration in the supernatants was measured with Bio-Rad Coumassie R250 reagent according to the protocols described by the producer. The chloramphenicol-acetyltransferase activity was measured according to the conventional method (Schottel JL, Sninsky JJ, Cohen SN "Effects of alterations in the translation control region on bacterial gene expression: use of cat gene constructs transcribed from the lac promoter as a model system." *Gene*, 28 (1984) 177-193). In these experiments the 5,5'-dithio-bis(2-nitrobenzoic acid) - the Ellman's reagent ("Sigma") was used as the specific reagent. The obtained results are shown in Table 1 of the main text.

SDS-PAGE (0.1% SDS - 12.5% PAAG electrophoresis) was performed for visualizing the accumulated CAT in the *E.coli* TG1-driven plasmid-carrier strains. The each strain was grown as described above. Each culture was divided into two parts. The cultures were supplemented with IPTG (up to 0.4 mM of the final concentration) and had been cultured for 2 hours. Then the cells were

harvested, resuspended in SDS-sample loading buffer (60 mM Tris-HCl pH6.8/2.3% SDS/10% glycerol/5% β -mercaptoethanol), and boiled for 15 minutes. 10-20 μ l of the resulting suspension as the sample was loaded on PAAG and electrophoration was performed according to the method described by Laemmli (Laemmli V.K.// "Cleavage of structural proteins during the assembly of the head of bacteriophage T4". Nature 227 (1970) 680-685). The gel was stained by Coomassie-blue for detecting the separated proteins. The corresponding patterns of the obtained gels are presented in the Figure 3.

SEQUENCE LISTING

<110> Ajinomoto Co., Ltd.

<120> Expression control sequence

<130>

<160> 18

<210> 1

<211> 118

<212> DNA

<213> Escherichia coli

<400> 1

atgaaagcaa ttttcgtact gaaaggttgg tggcgcaactt cctgaaacgg gcagtgtatt	60
caccatgcgt aaagcaatca gatacccagc ccgcctaatag agcgggcttt tttttgaa	118

<210> 2

<211> 153

<212> DNA

<213> Escherichia coli

<400> 2

atgacacgcg ttcaatttaa acaccaccat catcaccatc atcctgacta gtcttttcagg	60
cgatgtgtgc tggaagacat tcagatcttc cagtgggtgca tgaacgcatg agaaagcccc	120
cgggaagatca ctttcgggg gcttttttat tgc	153

<210> 3

<211> 169

<212> DNA

<213> Artificial sequence

<220>

<223> anti-attenuator

<400> 3

atgaaagcaa ttttcgtact gaaaggttgg tggcgcaactt cctgaaacgg gcagtgtatt	60
--	----

caccatgcgt aaagcaatca gatactagat ctgcccgact gcgtacaacg gtacagaaag 120
 cccccggcag atcacctgcc gggggctttt ttattggcgg ttgataacg 169

<210> 4

<211> 69

<212> DNA

<213> *Escherichia coli*

<400> 4

ctagaaagct taacacagaa aaaagccgc acctgacagt gcgggctttt ttttcgacc 60
 actgcagga 69

<210> 5

<211> 69

<212> DNA

<213> *Escherichia coli*

<400> 5

gatccctgca gtggtcgaaa aaaaaagccc gcactgtcag gtgcgggctt ttttctgtgt 60
 taagcttta 69

<210> 6

<211> 32

<212> DNA

<213> Artificial sequence

<220>

<223> primer

<400> 6

gcttaggtac cctcccatc cccctgttga ca 32

<210> 7

<211> 35

<212> DNA

<213> Artificial sequence

<220>

<223> primer

<400> 7

ctgtttctag atcctgtgtg aaattgttat ccgca 35

<210> 8

<211> 36

<212> DNA

<213> Artificial sequence

<220>

<223> primer

<400> 8

cagagctcta gaagttcacg taaaaagggt atcgac 36

<210> 9

<211> 35

<212> DNA

<213> Artificial sequence

<220>

<223> primer

<400> 9

gtatcgcata tgaaagcaat tttcgtactg aaagg 35

<210> 10

<211> 37

<212> DNA

<213> Artificial sequence

<220>

<223> primer

<400> 10

gtctgagatc tagtatctga ttgctttacg catgggt 37

<210> 11

<211> 42

<212> DNA

<213> Artificial sequence

<220>

<223> primer

<400> 11

atcataggat cctaattttg ttcaaaaaaa agcccgctca tt

42

<210> 12

<211> 47

<212> DNA

<213> Artificial sequence

<220>

<223> primer

<400> 12

cgactgtcta gaacggtaca gaaagccccc ggcagat

47

<210> 13

<211> 27

<212> DNA

<213> Artificial sequence

<220>

<223> primer

<400> 13

agctcaccgt ctttcattgc catacgg

27

<210> 14

<211> 32

<212> DNA

<213> Artificial sequence

<220>

<223> primer

<400> 14

acatgcggtta ccgatcccg c gaaattaata cg

32

<210> 15

<211> 65

<212> DNA

<213> Artificial sequence

<220>

<223> anti-attenuator fragment

<400> 15

cagagctcta gaagatctgc ccgactgcgt acaacggtac agaaagcccc cggcagatca 60
cctgc 65

<210> 16

<211> 43

<212> DNA

<213> Artificial sequence

<220>

<223> anti-attenuator fragment

<400> 16

cgggggcttt tttattgcgc ggttgataac gggatccagc gta 43

<210> 17

<211> 65

<212> DNA

<213> Artificial sequence

<220>

<223> anti-attenuator fragment

<400> 17

tacgctggat cccgttatca accgcgcaat aaaaaagccc ccggcaggtg atctgccggg 60
ggctt 65

<210> 18

<211> 43

<212> DNA

<213> Artificial sequence

<220>

<223> anti-attenuator fragment

<400> 18

tctgtaccgt tgtacgcagt cgggcagatc ttctagagct ctg

What is claimed is:

1. An expression control sequence which controls expression of a target gene linked downstream of the expression control sequence depending on an
5 intracellular concentration of an amino acid,

wherein in a bacterium which harbors a DNA construct comprising the expression control sequence, a promoter linked upstream of the expression control sequence and the target gene linked downstream of the
10 expression control sequence, frequency of termination in the expression control sequence, of transcription starting from the promoter is lowered by increase of an intracellular concentration of an amino acid, whereby expression of the target gene increases.

15 2. The expression control sequence according to claim 1, which comprises a region coding for a leader peptide comprising said amino acid and a ρ independent terminator, wherein when translation of the leader peptide stops at codon of said amino acid in the
20 course of the translation in case of starvation of said amino acid, a base pairing structure of the ρ independent terminator is formed in a transcript of the expression control sequence, whereby the frequency of termination in the expression control sequence, of the transcription
25 increases.

3. The expression control sequence according to claim 2, which comprises an odd number of not less than

3, of segments, wherein each of the segments can form a base paring structure together with its adjacent segment, and wherein in the transcript of the expression control sequence, when a segment or segments other than terminal segments each form a base paring structure with one of its two adjacent segments, the segment or segments each do not form a base paring structure with the other of the two adjacent segments; a first segment at an upstream terminal overlaps with the region interacting with the ribosome translating the leader peptide; a second segment adjacent to the first segment forms a base paring structure with a third segment adjacent to the second segment in the course of the translation of the leader peptide; and a base paring structure formed from the downstream terminal segment and its adjacent segment is the base paring structure of the ρ -independent terminator.

4. The expression control sequence according to claim 3, wherein the first segment overlaps with codon of the amino acid in the leader peptide.

5. The expression control sequence according to claim 3 or 4, wherein the number of the segments is 5.

6. The expression control sequence according to any of claims 3 to 5, wherein the sequence of each segment or a part thereof and the sequence of the adjacent segment or a part thereof constitute an inverted repeat sequence.

7. The expression control sequence according to any of claims 2 to 6, wherein the ρ -independent terminator is capable of functioning in a bacterium belonging to the genus *Escherichia*.

5 8. The expression control sequence according to claim 7, which comprises five segments an1 to an5 in order from an upstream side, wherein the segments an1 and an2, and a coding region for the leader peptide are derived from a sequence of an attenuator of a tryptophan
10 operon of *Escherichia coli*, the segments an4 and an5 are derived from a sequence of an attenuator of a histidine operon of *Escherichia coli*, and the segment an3 is derived from a combination of the sequences of the attenuators of the tryptophan operon and the histidine
15 operon.

9. The expression control sequence according to claim 8, the leader peptide has been modified to contain not less than 2 of tryptophan residues.

20 10. A method for controlling an expression of a target gene, comprising the steps of:

cultivating a bacterium harboring a DNA construct comprising the expression control sequence as defined in any of claims 1 to 9, a promoter linked upstream of the expression control sequence and the target gene linked
25 downstream of the expression control sequence in a culture medium, and

changing an intracellular concentration of an

amino acid on which expression control by the expression control sequence depend, to control expression of the target gene.

11. A method for producing a target substance
5 comprising the steps of cultivating a bacterium capable of producing the substance to produce the substance and collecting the substance,

wherein the bacterium harbors a DNA construct comprising the expression control sequence as defined in
10 any of claims 1 to 9, a promoter linked upstream of the expression sequence and a target gene which has relationship to production of the target substance and is linked downstream of the expression control sequence, and an intracellular concentration of an amino acid on
15 which expression control by the expression control sequence depend, is changed to control expression of the target gene.

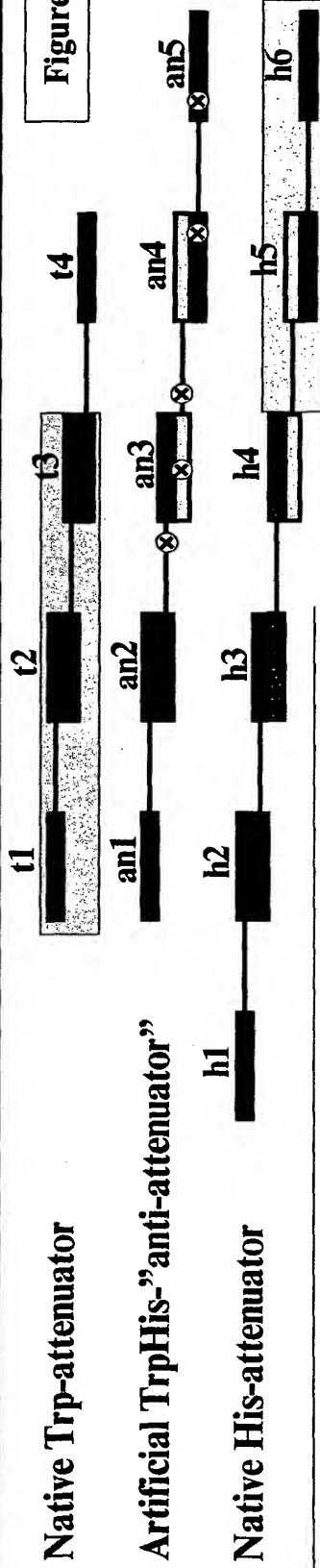
12. The method according to claim 11, wherein
the intracellular concentration of the amino acid is
20 changed by synthesis or degradation of the amino acid by the bacterium.

ABSTRACT OF THE DISCLOSURE

An expression control sequence which controls expression of a target gene linked downstream of the expression control sequence depending on an
5 intracellular concentration of an amino acid,
wherein in a bacterial cell which harbors a DNA construct comprising the expression control sequence, a promoter linked upstream of the expression control
sequence and the target gene linked downstream of the
10 expression control sequence, frequency of termination in
the expression control sequence, of transcription
starting from the promoter is lowered by increase of an
intracellular concentration of an amino acid, whereby
expression of the target gene increases.

THE STRUCTURE AND PROPOSED PROPERTIES OF THE NEW ARTIFICIAL "ANTI-ATTENUATOR" REGULATORY ELEMENT.

Figure 1.



⊗ -modification

RNA polymerase stops at pause site in attenuator region [J.Mol.Biol.(1997),268,54-68]

Ribosome binds, starts translation and releases the RNA polymerase



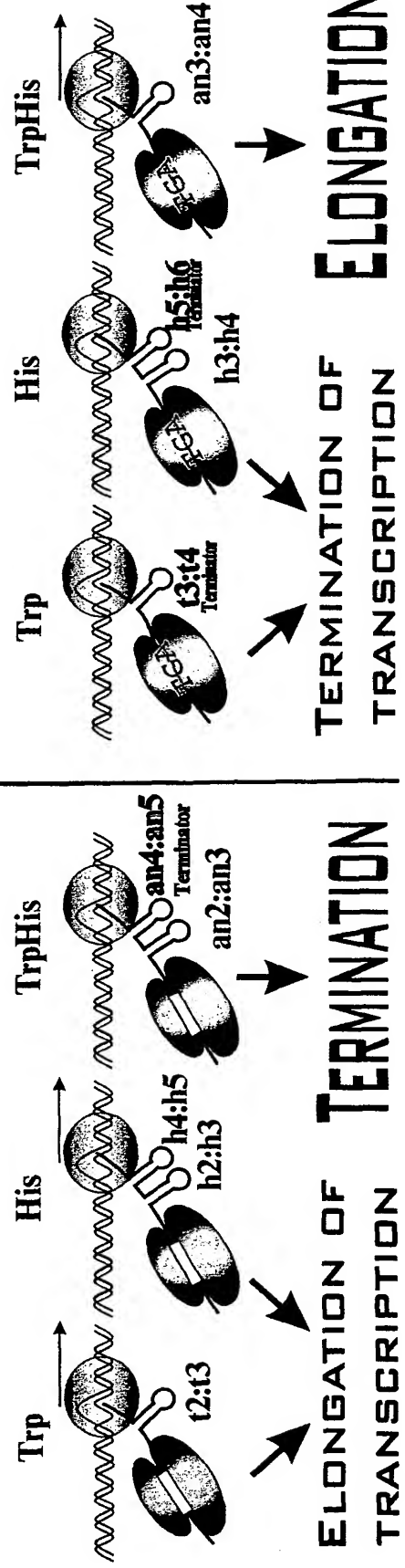
STARVATION

Ribosome stalls at "sense" aminoacid codons

"SENSE" AMINOACID

EXCESS

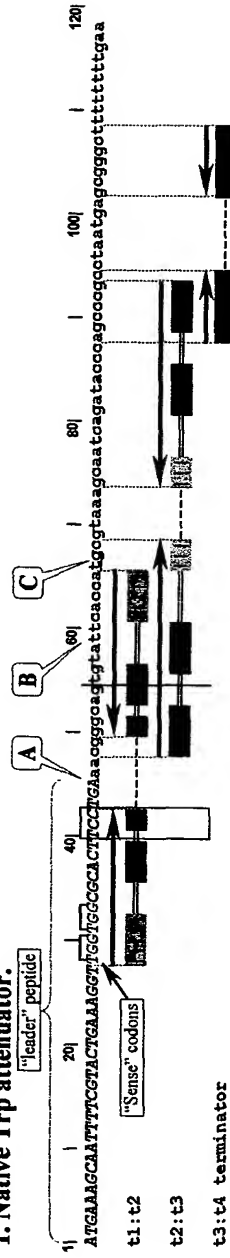
Ribosome stalls at STOP codon



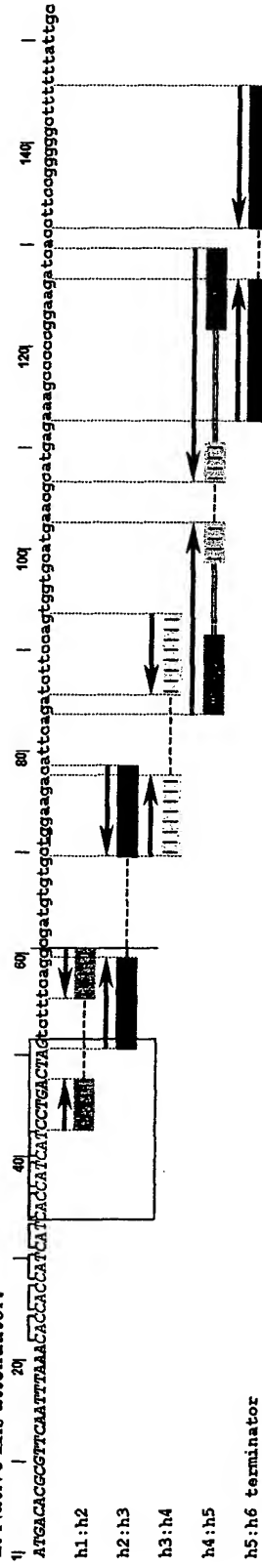
DETAILED STRUCTURES OF THE NATIVE ATTENUATORS AND THE ARTIFICIAL ANTI-ATTENUATOR

Figure 2

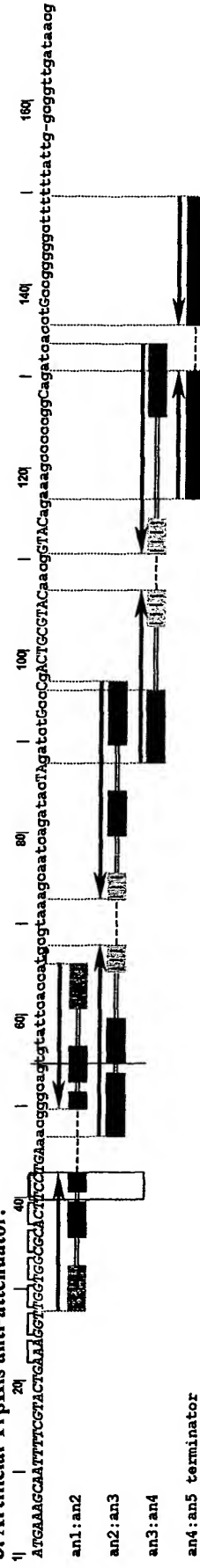
1. Native Trp attenuator.



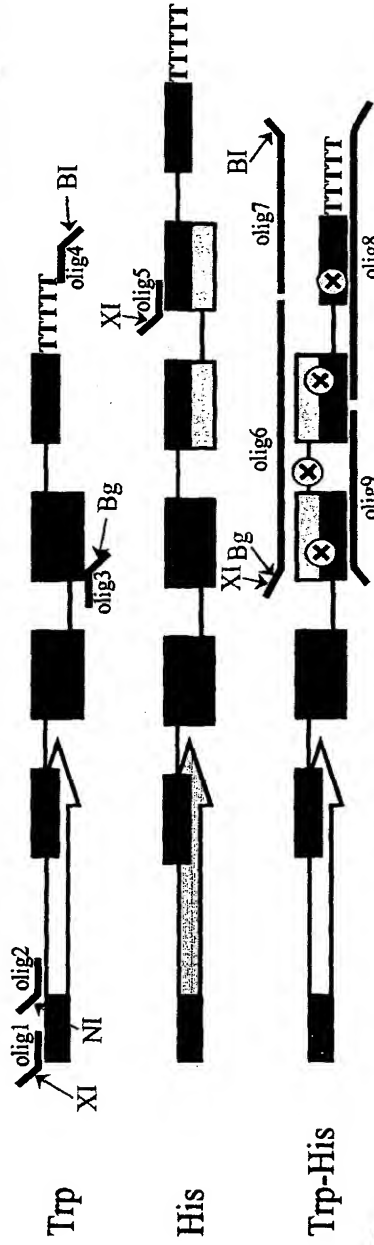
2. Native His attenuator.



3. Artificial TrpHis anti-attenuator.



General Scheme of the New Artificial "Anti-Attenuator" Regulatory Element Construction.



⊗ - nucleotide changing in comparison to the native sequence.

Fragment	Procedure	Resulting Construction	Plasmid Name	Properties
1	<ul style="list-style-type: none"> Phosphorylation of olig7 and olig9 Annealing & ligation ↓XI & ligation with V↓XI ↓BI & cyclization 		pML-P _{lac} -an3-an4(an4-an5)-cat	1 Antiterminator
2	<ul style="list-style-type: none"> ↓XI Ligation with V↓XI ↓BI & cyclization 		pML-P _{lac} -an4-an5-cat	2 Terminator
3	<ul style="list-style-type: none"> ↓XI Ligation with V↓XI ↓BI & cyclization 		pML-P _{lac} -trpL-cat	3 Trp-attenuator
4	<ul style="list-style-type: none"> ↓XI Ligation with V↓XI ↓Bg & cyclization 		pML-P _{lac} -anti-att-I-cat	4 Antiattenuator 1
5	<ul style="list-style-type: none"> ↓XI Ligation with V↓XI ↓Bg & cyclization 		pML-P _{lac} -anti-att-II-cat	5 Antiattenuator 2

Figure 3.

BI-BamHI, Bg-BglII, NI-NdeI, XI-XbaI, XI*-XbaI(claim-)

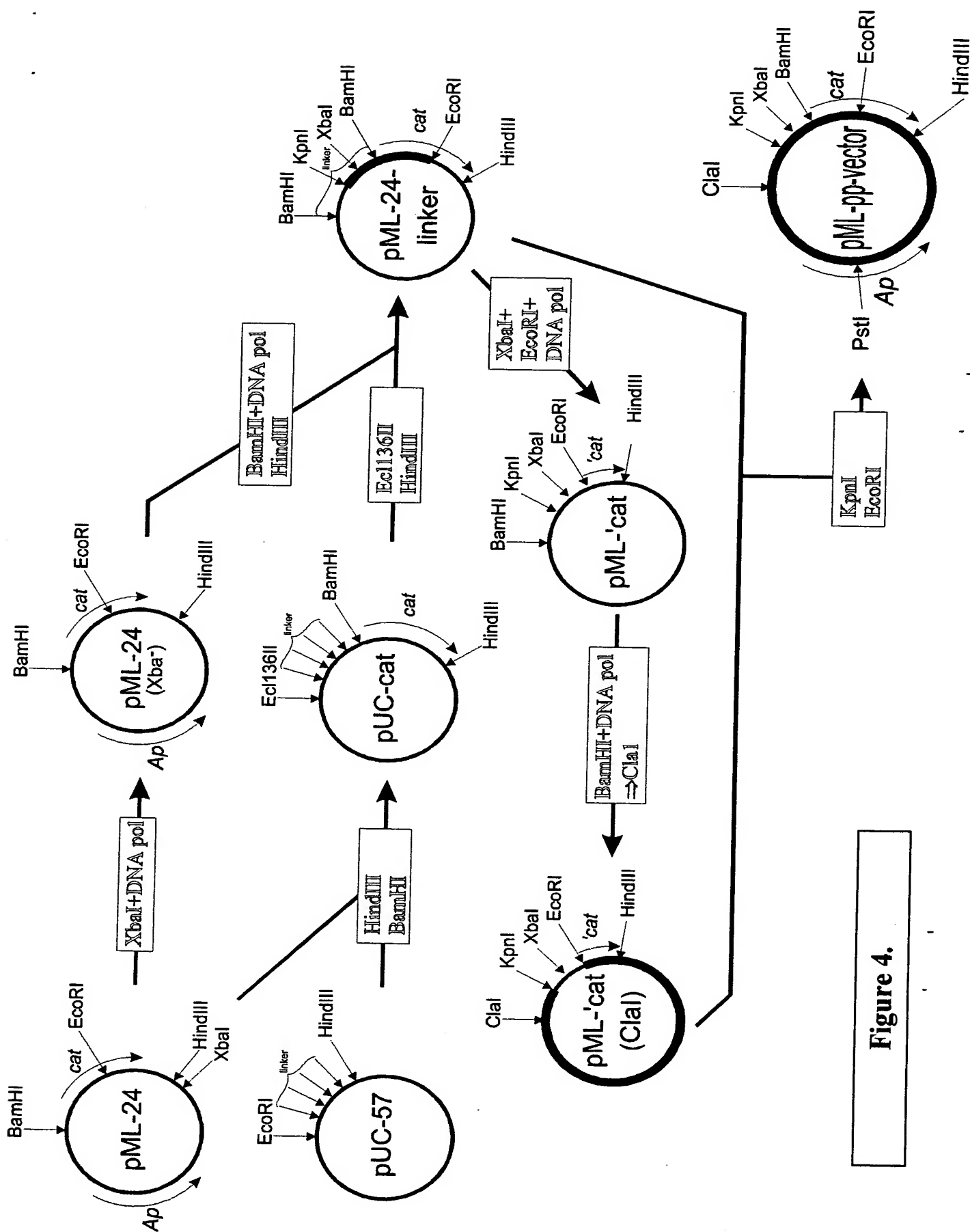


Figure 4.

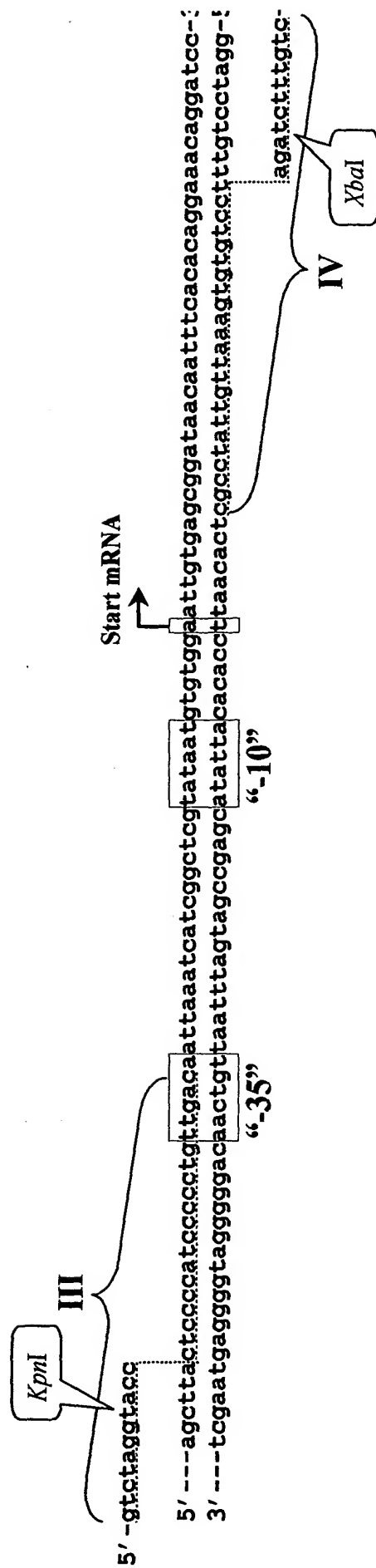


Figure 5. Structure of the P_{lac} promoter from the plasmid pDR540. The sequences of the "upstream" (III) and "downstream" (IV) primers for the PCR are dot-matched.

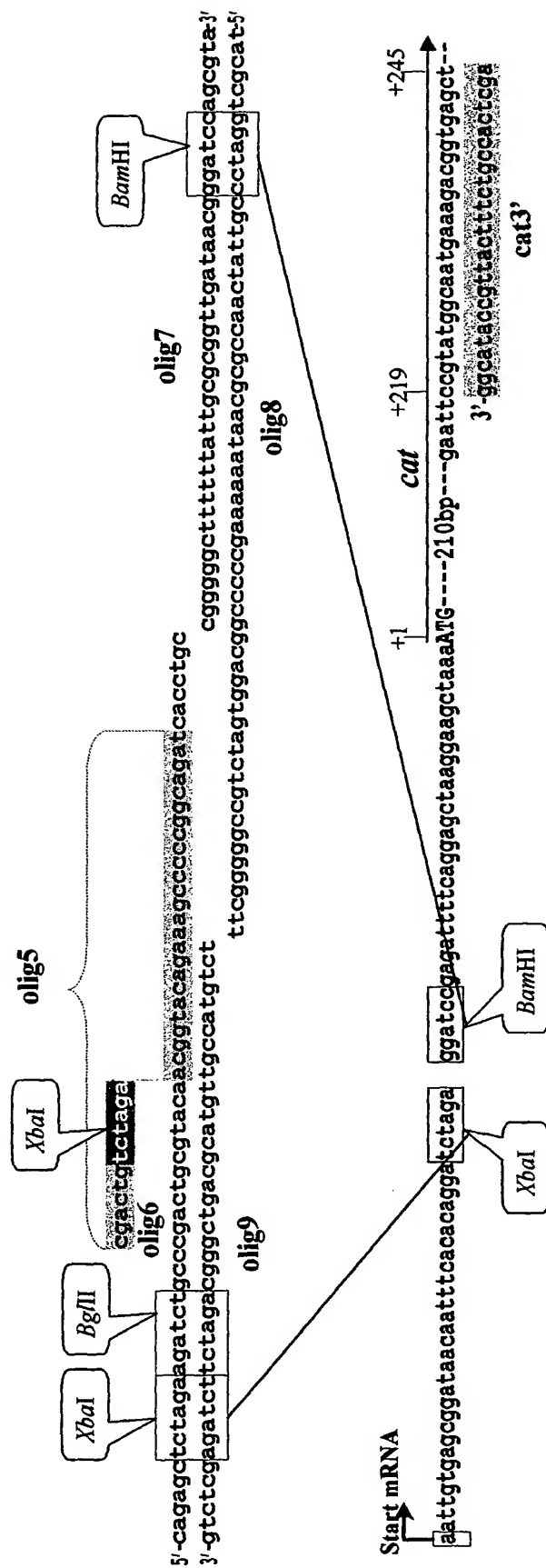


Figure 6. The structure of chemically synthesized an3:an4(an4:an5) fragment and the way it was inserted into the cloning vector.

